AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 139:16–22 (2009)

Race and Global Patterns of Phenotypic Variation

John H. Relethford*

Department of Anthropology, State University of New York College at Oneonta, Oneonta, NY 13820

KEY WORDS race; skin color; craniometrics

ABSTRACT Phenotypic traits have been used for centuries for the purpose of racial classification. Developments in quantitative population genetics have allowed global comparison of patterns of phenotypic variation with patterns of variation in classical genetic markers and DNA markers. Human skin color shows a high degree of variation among geographic regions, typical of traits that show extensive natural selection. Even given this high level of geographic differentiation, skin color variation is clinal and is not well described by discrete racial categories. Craniometric traits show a level of among-region differentiation comparable to genetic markers, with high levels of variation within populations as well as a correlation between phenotypic and geographic distance. Craniometric variation is geographically structured, allowing high levels of classification accuracy when comparing crania from different parts of the world. Nonetheless, the boundaries in global variation are not abrupt and do not fit a strict view of the race concept; the number of races and the cutoffs used to define them are arbitrary. The race concept is at best a crude first-order approximation to the geographically structured phenotypic variation in the human species. Am J Phys Anthropol 139:16–22, 2009. © 2009 Wiley-Liss, Inc.

The question of biological race as a tool for describing human biological variation has a long history of debate in physical anthropology. No attempt is made here to review this history in detail, something covered in detail in a number of books (e.g., Kennedy, 1976; Shanklin, 1994; Marks, 1995; Graves, 2001; Brace, 2005, among others). It is clear, however, that there is still debate over the utility of biological concepts of race in describing human variation, with some arguing for its use (e.g., Sarich and Miele, 2004) and others arguing against it (e.g., Brace, 2005). There is also debate over the extent to which support of the race concept within the field of physical anthropology has changed over time (Cartmill and Brown, 2003; Lieberman et al., 2003).

The term "race" has been used to refer to the aspects of both biological and cultural variation and has been applied to everything from geography to genes and from ancestry to language. Even if we restrict the usage to biological race, there is still debate on how to apply the term, and even if the term should be applied at all. Most current uses of the term are generally geographically based, such as the idea of geographic race articulated by Garn (1965), which generally subdivides humanity into continental groupings (e.g., Africans, Asians, Europeans) and/or subcontinental groupings (e.g., East Asians, Southeast Asians).

The use of geographically defined races is common across many disciplines, but this does not mean that it is the best way to describe or analyze human biological variation. We can all agree that human biological variation exists. The important question concerns the best way of describing and analyzing this variation, and how well this goal is met by the race concept. Much current discussion over race revolves around the use of a rapidly growing set of genetic markers and their ability for classification (e.g., Rosenberg et al., 2002; Bamshad et al., 2004; Jorde and Wooding, 2004), whereas most classic definitions of race focus on phenotypic traits, such as skin color, craniofacial shape, and hair color and texture, among other features. Certainly, these phenotypic traits are what people use for classification in everyday contexts; one sees physical variations, such as skin color, when walking down the street, and not DNA haplotypes. Given the historic and contemporary use of phenotypic traits in racial classification, it is appropriate to examine their global distribution to determine if, and how well, a racial approach fits the observed data. The purpose of this article is to examine the global distribution of two examples of phenotypic traits, skin color and craniometric traits, with an eye toward the utility of the race concept. Previous studies focusing on quantitative traits from a population-genetic perspective are reviewed here and their results considered in light of the question of race and racial classification. To meet the stated needs of this symposium, I conclude the article with some general thoughts regarding the general question of the usefulness of the biological race concept. A number of these concerns have long been noted in the literature on subspecies in zoological classification (e.g., Wilson and Brown, 1953; Futuyma, 1986).

GLOBAL VARIATION IN SKIN COLOR

Skin color has played a role in virtually all racial classifications to date. Although skin color is often described in the public sphere using discrete labels, such as "black," "brown," and "white," the actual distribution of human skin color is continuous and cannot be described

*Correspondence to: John H. Relethford, Department of Anthropology, SUNY College at Oneonta, Oneonta, NY 13820, USA. E-mail: relethjh@oneonta.edu

Received 21 March 2008; accepted 17 June 2008

DOI 10.1002/ajpa.20900

Published online 18 February 2009 in Wiley InterScience (www.interscience.wiley.com).

© 2009 WILEY-LISS, INC.

in terms of a finite number of shades. Anthropologists have measured skin color using one of several models of reflectance spectrophotometers, devices that measures the amount of light at a given wavelength that is reflected back off a person's skin; the lighter the skin, the greater the amount of reflected light. The majority of studies of human skin color to date have relied on assessing skin color using filter number 609 of the E.E.L. reflectometer; this filter samples the visible wavelength at 685 nm, a wavelength that correlates with melanin concentration (Relethford, 1997, 2000).

The first question regarding global variation in human skin color is the degree of variation among continental and subcontinental groups. Population geneticists often express the degree of variation among populations using a measure known as Wright's F_{ST}, which is the amount of genetic variation among groups relative to a panmictic state. F_{ST} is often expressed as a percentage; for example, an F_{ST} of 12% means that 12% of the genetic variation exists among groups and consequently that 88% of the genetic variation exists within groups. The use of F_{ST} in critiques of racial classifications dates back to Lewontin's (1972) pioneering work on the global distribution of classical genetic markers. Some have argued that the relatively low F_{ST} found in studies of human genetic markers is inconsistent with a racial approach that stresses discrete and "sharply differentiated populations" (Templeton, 1998: 633). Although the usefulness of some formulations of F_{ST} for assessing the utility of the race concept has been challenged (e.g., Long and Kittles, 2003; Sarich and Miele, 2004), it remains a useful statistic for comparing the overall levels of regional differentiation.

F_{ST} can be estimated from phenotypic traits given heritability estimates (Williams-Blangero and Blangero, 1989; Relethford and Blangero, 1990; Relethford et al., 1997). Elsewhere (Relethford, 2002), I used data on skin color (percent reflectance for E.E.L. filter 609) from 68 samples of males to estimate F_{ST} from skin color. Using a heritability of 0.66 for filter 609 (Williams-Blangero and Blangero, 1992), I estimated F_{ST} to be 0.88 when subdividing the total sample into five geographic regions: sub-Saharan Africa, Europe, Central Asia, Australasia, and the Americas. When the analysis is confined to the three geographic regions often used in everyday concepts of geographic race (sub-Saharan Africa, Europe, and Asia), the F_{ST} estimate is virtually the same (0.87). Even given caveats regarding heritability estimation, the number and composition of regions, and different computational methods used to estimate F_{ST}, the overall result shows clearly that skin color variation among geographic regions is far above that found for classical genetic polymorphisms and DNA markers, most of which typically show values of F_{ST} among regions of about 0.10–0.15 (Barbujani et al., 1997; Relethford, 2002). Regardless of whether one considers the F_{ST} values from genetic markers useful or not useful when assessing the race concept, these comparisons show that skin color, a phenotypic trait that has been central to most racial classification schemes, actually has an atypical pattern of among-group diversity. Nevertheless, the elevated degree of differentiation in human skin color by itself does not provide support for a view of discrete and differentiated human races.

Elevated values of F_{ST} across geographic regions are often associated with natural selection inflating the amount of among-group variation, such as the Duffy negative allele (Cavalli-Sforza et al., 1994). Natural selection also explains the high F_{ST} for skin color. A number of studies have shown a very strong relationship between skin color and latitude and between skin color and the intensity of ultraviolet radiation (e.g., Roberts and Kahlon, 1976; Relethford, 1997; Jablonski and Chaplin, 2000).

Even though regional variation in skin color does fit a common perception regarding high levels of differentiation among geographic races, the question remains if there are any clusters in skin color that correspond to different views on discrete races. Everyday discussion of skin color often tends to focus on a small number of discrete groups corresponding with terms such as "white," "black," and "brown." Given the association of skin color and latitude, it is easy to separate people from widely different geographic regions into different groups based on skin color. If, for example, we took skin reflectance measures on natives from a village in Scotland and from a village in Ethiopia, it would be trivially easy to separate the two samples with 100% accuracy. This does not mean, however, that skin color supports the idea that there are two discrete races. This would only be the case if everyone in the world were a native of the village in Scotland or the village in Ethiopia. This is of course not the case; many other populations bridge the gap between these two, and the association of skin color and latitude does not provide easily identifiable clusters.

The continuous nature of human skin color variation is shown in Figure 1a, which plots the mean $\pm$ 1 SD, for male skin color in 98 samples from the Old World for which standard deviations were available (Relethford, 2000). The distribution of each sample is plotted on a separate line to allow the comparison of the distribution of skin color in different populations. Close inspection of these values shows that each population overlaps with one or more other populations. Figure 1b shows what the worldwide distribution in skin color looks like when all of these separate distributions are overlain on a single line; the result is a continuous straight line ranging from the darkest extremes to the lightest extremes in skin color. There are no identifiable clusters. It is clear that people from the left side of the graph are dark-skinned and people from the right side are light-skinned, but it is also clear that every value in between is represented. Researchers are of course free to subdivide this continuum into different groups, but such clustering would be arbitrary and subjective in terms of the number of groups and the cutoff points used to distinguish them. The lack of apparent clusters is a reflection of the fact that skin color shows a classic pattern of clinal variation.

GLOBAL VARIATION IN CRANIOMETRICS

Not all phenotypic traits show the extreme global differentiation and environmental correlation of skin color. Here, I focus on global patterns of craniometric variation. Craniometric traits have been shown to be useful in reconstructing population history and in assessing ancestry. Not all craniometrics are selectively neutral. Although some traits show correlations with climate, previous studies have shown that the signal from population history is not erased by these trends (e.g., Relethford, 2004a; Roseman, 2004). The average patterns of global craniometric variation often provide a reasonable reconstruction of population history.

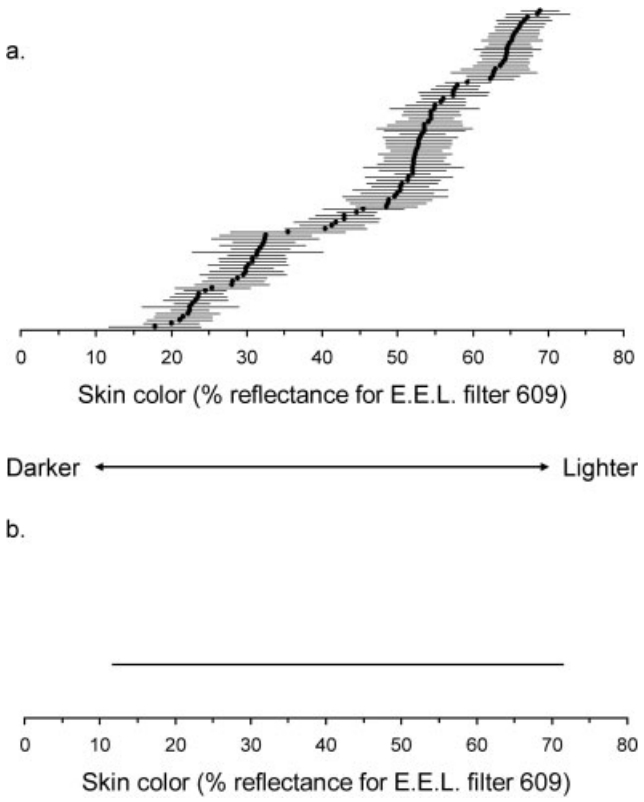

Fig. 1. Distribution of skin color in 98 Old World male samples as assessed using filter 609 for the E.E.L. portable reflectance spectrophotometer. Filter 609 samples the visible wavelength at 685 nm. Data sources are given in Relethford (2000). (**a**) The distribution of each population is represented using a dot for the mean and the lines on each side of the mean to indicate ± 1 SD. Populations are arranged vertically and sorted from darkest to lightest to show overlap between populations. (**b**) All 98 distributions are combined on a single axis to show total overlap with no discrete breaks in skin color. Human skin color does not come in a finite number of shades but instead ranges continuously from dark to light.

TABLE 1. Sample sizes and location of Howells's 18 core populations

Region	Population	Sample size	
		Males	Females
Europe	Norse, Norway	55	55
	Zalavar, Hungary	53	45
	Berg, Austria	56	43
Sub-Saharan Africa	Teita, Kenya	33	50
	Dogon, Mali	47	52
	Zulu, South Africa	55	46
Australasia	Australia	52	49
	Tasmania	45	42
	Tolai, New Britain	56	54
Polynesia	Mokapu, Hawaii	51	49
	Easter Island	49	37
	Moriori, Chatham Islands	57	51
Americas	Arikara, South Dakota	42	27
	Santa Cruz, California	51	51
	Peru (Yauyos District)	55	55
East Asia	North Japan	55	42
	South Japan	50	41
	Hainan, China	45	38

Several studies have looked at estimates of F_{ST} based on the global craniometric dataset originally collected by Howells (1973, 1989, 1995, 1996), and consisting of 57 craniometric variables measured on male and female crania from 26 populations (as well as males only on four additional samples). Using an average heritability of 0.55, Relethford (1994, 2002) found that estimates of F_{ST} based on all 57 traits ranged from 0.11 to 0.14 depending on the number of geographic regions sampled. These F_{ST} values are similar in magnitude to those estimated in a number of studies of classical genetic markers and DNA markers. Further, subdivision of variation within the geographic regions shows that most of the variation is within local populations, again corresponding to results from genetic marker studies (Relethford, 2002). Roseman and Weaver (2004) investigated variation in F_{ST} estimates using principal components based on Mosimann shape variables and found that some craniometric traits have higher F_{ST} values, suggesting some regional differentiation due to selection.

Measures of F_{ST} focus on the degree of variation, whereas there are a number of questions relating to the race concept that deal with the pattern of variation. Does craniometric variation produce recognizable clusters? What is the relationship of craniometric variation to geography and neutral genetic markers? Can we accurately classify crania, and if so, then what does that tell us about the utility of the race concept? Some of these questions have been answered elsewhere in a different context. For the purposes of this article, I focus here on analyses using data on Howells's "core" sample-crania on 907 males and 827 females from three local populations in each of six geographic regions: Europe, sub-Saharan Africa, Australasia, Polynesia, the Americas, and East Asia. The location and sample sizes of these 18 samples are reported in Table 1.

The first question relating to race is whether craniometric variation forms recognizable clusters. To answer this, I computed a phenotypic distance matrix between the 18 populations using methods outlined elsewhere (Williams-Blangero and Blangero, 1989; Relethford and Blangero, 1990; Relethford et al., 1997). These distances were derived using a pooled sample of all male and female crania after standardizing all values within each sex (Relethford, 1994). Figure 2 shows the results of an UPGMA (Unweighted Pair Group Method with Arithmetic Mean) cluster analysis on the distance matrix. Although the clusters tend to be long, there is clear separation into the six geographic regions sampled by Howells. These clusters reflect the fact that craniometric variation (like genetic marker and DNA marker variation) is geographically structured and that Howells sampled regions at far distances from each other rather than a random geographic sampling.

The geographic structure of craniometric variation is clearer when examining the relationship between phenotypic distance and geographic distance between all pairs of populations. To investigate this relationship, I derived geographic distances between all pairs of populations. As described elsewhere (Relethford, 2004a,b), geographic distances were obtained after constraining interpair distances to likely migration routes between geographic regions: through the Bering Strait for distances to New World populations, through Northeast Africa for populations in Africa, and through Southeast Asia for populations in Australasia and the Pacific. The plot of phenotypic distance versus geographic distance is shown in Figure 3a, which shows clearly that populations farther

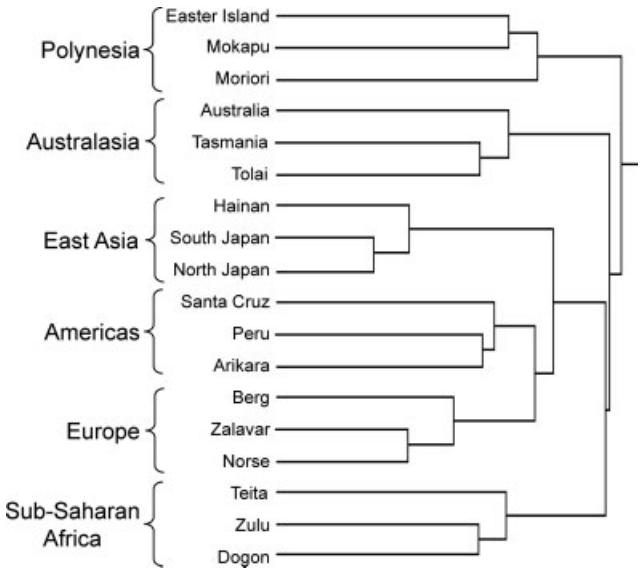

Fig. 2. Dendrogram showing the phenotypic distance between Howells's 18 core populations. The UPGMA clustering algorithm was used after taking the square root of the Relethford-Blangero distances (which are squared distances).

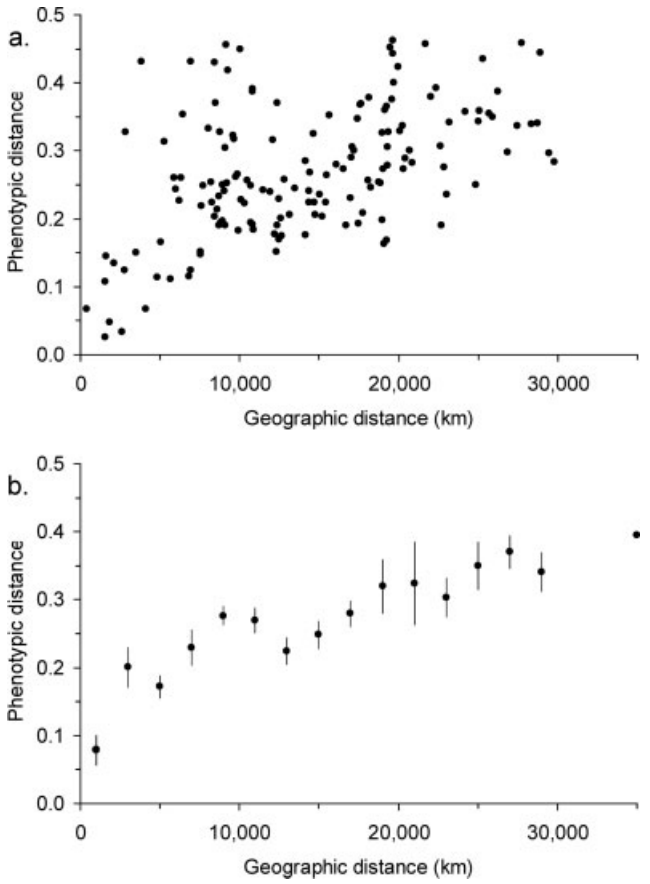

Fig. 3. Relationship between phenotypic distance and geographic distance between all pairs of populations using Howells's 18 core populations. (**a**) Plot of each individual distance pair. (**b**) Distances pooled by geographic distance using 1,000 km intervals up to 20,000 km and 2,000 km intervals afterward. Dots represent the mean phenotypic distance in each geographic distance interval and vertical lines represent the means ±1 standard error.

apart geographically tend to be more dissimilar phenotypically (a correlation of $r = 0.521$, $P = 0.002$, based on 9,999 random permutations using the Mantel test). The association of phenotypic and geographic distance is even more apparent in Figure 3b, which plots the mean distances for different intervals of geographic distance.

Comparative studies show that classical genetic markers, microsatellite DNA markers, and craniometrics all fit this pattern of a global relationship between genetic/phenotypic similarity and geography (Relethford, 2004b). In addition, global patterns of craniometric variation and microsatellite DNA variation produce similar patterns of population relationship (Roseman, 2004). These studies all show that global patterns of craniometric variation reflect largely underlying patterns of genetic relationship, which in turn reflect geographic structure. There is some debate over the cause(s) of the genetics-geography correlation. Is it simply the global extension of the isolation by distance model, whereby geographically proximate populations are more likely to share genes than with populations farther away? Alternatively, is it a reflection of the geographic dispersion of modern humans out of Africa through a series of founding events (Ramachandran et al., 2005, but see Templeton, 2007)? Of course, these possibilities are not mutually exclusive (Relethford, 2004b). This debate is bound to continue but is not directly relevant here to the issue of racial classification.

What is relevant is that there is an association of craniometric variation and geography. This means that populations that are farther apart from one another will tend to be less similar than those populations that are closer to each other. Thus, when Howells sampled populations from six different geographic regions, he sampled populations that would be phenotypically different to begin with because of the correlation between geographic and phenotypic distance. There are no abrupt breaks in the relationship between phenotypic and geographic distance in Figure 3, indicating that decisions for subdivision into clusters (or races) are going to be subjective. Geographic race is a crude way of trying to describe the geographic structuring of human biological variation.

CLASSIFICATION USING CRANIOMETRICS

I now turn to the issue of racial classification in order to consider the geographically structured variation in craniometrics in a different light. Arguments for and against application of the race concept to humanity have often focused on the ability to accurately classify individuals into different racial groupings. I approach this question first by using discriminant analysis to determine if the cranial measures taken by Howells can be used to classify accurately crania into their respective geographic regions. Using Howells's six geographic regions and all 57 measurements, the overall rate of correct classification is 97% for male crania and 96% for female crania. Jackknifed classification rates are slightly lower but still very high (96% for males, 94% for females). Such high levels of classification would seem to suggest that dividing humanity (or at least that section of humanity represented by Howells's samples) into six discrete regions is valid. To argue against this subdivision in the face of such high accuracy often seems counterintuitive . . . at first.

TABLE 2. Classification accuracy of discriminant analysis using different regional groupings

Number of regions	Regions	Classification accuracy (%)[a]	
		Males	Females
6	Howells's core regions: Europe, sub-Saharan Africa, Australasia, Polynesia, Americas, East Asia	97 (95–99)	96 (89–99)
5	Howells's core regions with Americas and East Asia pooled	97 (96–99)	97 (94–99)
7	Howells's core regions with Americas divided into two regions: North America and South America	97 (95–99)	97 (90–99)
7	Howells's core regions with sub-Saharan Africa divided into two regions: Central Africa and South Africa	97 (93–99)	96 (89–99)
8	Howell's core regions with both Americas and sub-Saharan Africa each divided into two regions as above	96 (93–100)	96 (90–99)

[a] Overall classification rates; numbers in parentheses indicate the range of classification accuracy for local populations in different regions.

The relationship between racial units and classification accuracy is something often discussed in terms of forensic anthropology, where race, along with sex and age, are common elements in skeletal identification. The problem of classification was noted by forensic anthropologist Norman Sauer, whose 1992 paper on forensic anthropology and race is subtitled with the important underlying question: "If races don't exist, why are forensic anthropologists so good at identifying them?" (Sauer, 1992). Sauer goes on to argue that classification accuracy does not validate the underlying taxonomic question of race:

> Some of the confusion about this issue may stem from an assumption that to identify a specimen as having ancestors in Africa or Europe, for example, is tantamount to race identification and a verification of geographic races. No one who argues against the race concept denies that human variation exists or claims that this variation is not systematic. In fact, it is systematic variation that allows anyone to estimate, with varying degrees of specificity, a person's place of ancestry from their physical features. However, to identify a person as having ancestors from, say, Northern Europe does not identify a biological race of Northern Europeans.
>
> —Sauer, 1992: p 110

I explore the issue of classification further by considering what happens to classification accuracy if the geographic regions used by Howells are defined in slightly different ways. That is, what happens if we pool regions into a smaller number, or subdivide them into a larger number? Here, I consider several alternatives to the six-region classification using Howells's set of 18 populations. One alternative is to lump the East Asian and Native American samples based on recent shared ancestry, creating a smaller five-region analysis. Other alternatives involved splitting into more than six regions, such as dividing the Native American sample into North American and South American regions ($n = 7$ regions), dividing the African sample into Central African and South African regions ($n = 7$ regions), or dividing both American and African samples ($n = 8$ regions). The results of these different scenarios are presented in Table 2; in each case, classification accuracy remained high and about the same as for the six-region analysis (96–97%). Jackknifed classification rates (not shown) are also similar. Table 2 also reports variation in the range of classification accuracy for individual regions, and these ranges are also roughly the same in each classification scenario. If defining different numbers of geographic races all produce essentially the same rate of accuracy in classification, then what justification exists for preferring one number of races to another? For example, should we group the Native American samples into one race or two?

Although I have not explored all possible permutations, I suggest that if one defines clusters in such a way that geographical distances among clusters typically exceed geographic distance within clusters, and if one uses enough traits in the analysis, then there will be a high degree of accuracy in classification. These results again stem from the simple fact that people living farther apart in the world tend to look more different combined with a large number of traits. The same results apply when classifying individuals into major geographic regions using a large number of DNA markers (e.g., Rosenberg et al., 2002). Given enough markers, one can easily assign individuals to geographically widespread groupings such as Western Europe, East Asia, and sub-Saharan Africa. These analyses are simply another way of expressing the correlation between genetics and geography. There is no justification based on classification accuracy for preferring five, six, seven, or eight groups. From the perspective of racial classification, if we take these different groups as equivalent to races, then we see that the number of races, and the geographic cutoffs used to define them, are subjective decisions.

DO RACES EXIST?

Does the global pattern of phenotypic variation seen in skin color and craniometric traits support the idea of race? The answer depends on initial definitions and use of the term "race." On the one hand, the idea of discrete races in the typological mindset of past centuries clearly does not apply to humans (Templeton, 1998; Graves, 2001). There is, however, geographic patterning to human biological variation, both for traits affected strongly by natural selection, such as skin color, and for traits whose distribution tends, on average, to be more neutral, such as craniometrics and many genetic polymorphisms. I suggest that typical uses of the concept of geographic race today are simply crude labels imposed upon this geographically structured variation. In that sense, race is culturally constructed, as all labels are, but it is also based on an underlying reality of biological variation. Rather than argue about whether race is a cultural construct (an idea that many unfortunately take as being equivalent to a denial of variation) or that race is "real," it might be more useful to consider race as a culturally constructed label that crudely and imprecisely describes real variation.

Elsewhere (Relethford, 2008), I use an analogy of human height to describe students the difference

between underlying reality and cultural labels. Adult height in any human population can be described in terms of a normal distribution, where height is a continuous variable ranging from the shortest to the tallest individual in the population. Even given this continuous variation, in everyday use we often use labels such as "short," "medium," and "tall" as crude labels that are imposed on the underlying continuous variation. The relationship between the labels and the underlying reality is accurate in the sense that someone who is "short" is by definition shorter from someone who is "tall." This relationship, however, is crude and imprecise. It is subjective in terms of the number of possible groups (e.g., should we add "medium-short" and "very-tall"?) and the cutoffs used to define these groupings. For example, what value separates a person of "medium" height from a "tall" person? 1,800 mm?, 1,810 mm?, or 1806.25 mm? My point is that we tend to use crude labels in everyday life with the realization that they are fuzzy and subjective. I doubt anyone thinks that terms such as "short," "medium," and "tall" refer to discrete groups, or that humanity only comes in three values of height!

If we consider these questions for the moment in purely statistical terms, then we can see that the cultural construction of race is the transformation of a continuous variable into an ordinal-level or nominal-level variable with the attendant loss of statistical information. We do this all the time in both everyday life and research, using groupings such as social class, high versus low blood pressure, and liberal versus conservative political ideologies, to name but a few examples. In all such cases, the underlying variation is transformed into a set of discrete categories. Political ideology in the United States is an excellent example of the problems of information loss when using discrete categories; the nature of past elections has lead to a view of the United States as being made up of "red states" and "blue states" even though the underlying variation is continuous (Fiorina et al., 2005).

In such cases, the issue is whether reduction of continuous variation into discrete groups is appropriate and under what circumstances. In terms of the geographic structure of human biological variation, there are times when a crude division into major geographic regions may be useful, whereas in other cases such groupings would obscure too much of the variation, and analysis would be better performed using the local population as the unit of analysis. In some of my research, the objectives of the study were satisfied by taking major geographic regions, such as Europe and sub-Saharan Africa, as the units of analysis (e.g., Relethford, 1994; Relethford and Harpending, 1994). In other cases, such as examining the specific relationship of phenotypic distance and geography, local populations make better units of analysis (e.g., Relethford, 2004a,b). Broad geographic groupings might be useful in certain cases, such as a variety of forensic contexts where one might simply want to assign a specimen to a broad ancestral group. In other cases, the situation might be different and using such broad groupings would obscure too much variation. An example here would be using broad groupings in genetic studies of disease in African Americans, where the actual level of African ancestry is highly variable (Parra et al., 2001), even though most racial classifications would place all individuals in the same category.

The final question is one of semantics. In cases where broad geographic groups are used, should we refer to these groups as "races" or should we use more politically correct terms such as "geographic regions" or "geographic clusters?" On the one hand, the very concept of "race" has such historical baggage that continued use of the term tends to reify incorrect conceptions of human variation. On the other hand, it is probably naïve to think that the term can be wiped from everyday use and misuse. It might be more appropriate to answer the questions about race by noting that race is a crude first-order approximation to human biological variation that is arbitrary in terms of the number and definition of races. As such, race may not provide the best way of describing or analyzing human variation.

LITERATURE CITED

Bamshad M, Wooding S, Salisbury BA, Stephens JC. 2004. Deconstructing the relationship between genetics and race. Nat Rev Genet 5:598–609.

Barbujani G, Magagni A, Minch E, Cavalli-Sforza LL. 1997. An apportionment of human DNA diversity. Proc Natl Acad Sci USA 94:4516–4519.

Brace CL. 2005. "Race" is a four-letter word. New York: Oxford University Press.

Cartmill M, Brown K. 2003. Surveying the race concept: a reply to Lieberman, Kirk, and Littlefield. Am Anthropol 105:114–115.

Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton: Princeton University Press.

Fiorina MP, Abrams SJ, Pope JC. 2005. Culture war? The myth of a polarized America, 2nd ed. New York: Longman.

Futuyma DJ. 1986. Evolutionary biology, 2nd ed. Sunderland, MA: Sinauer Associates.

Garn SM. 1965. Human races, 2nd ed. Springfield, IL: Charles C Thomas.

Graves JL Jr. 2001. The emperor's new clothes: biological theories of race at the millennium. New Brunswick, NJ: Rutgers University Press.

Howells WW. 1973. Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human populations. Papers of the Peabody Museum No. 67. Cambridge, MA: Peabody Museum.

Howells WW. 1989. Skull shapes and the map: craniometric analyses in the dispersion of modern *Homo*. Papers of the Peabody Museum No. 79. Cambridge, MA: Peabody Museum.

Howells WW. 1995. Who's who in skulls: ethnic identification of crania from measurements. Papers of the Peabody Museum No. 82. Cambridge, MA: Peabody Museum.

Howells WW. 1996. Howells' craniometric data on the internet. Am J Phys Anthropol 101:441–442.

Jablonski NG, Chaplin G. 2000. The evolution of human skin coloration. J Hum Evol 39:57–106.

Jorde LB, Wooding SP. 2004. Genetic variation, classification and 'race.' Nat Genet 36:S28–S33.

Kennedy KAR. 1976. Human variation in space and time. Dubuque, IA: Wm. C. Brown.

Lewontin RC. 1972. The apportionment of human diversity. Evol Biol 6:381–398.

Lieberman L, Kirk RC, Littlefield A. 2003. Perishing paradigm: race-1931–99. Am Anthropol 105:110–113.

Long JC, Kittles RA. 2003. Human genetic diversity and the nonexistence of biological races. Hum Biol 75:449–471.

Marks J. 1995. Human biodiversity: genes, race, and history. New York: Aldine de Gruyter.

Parra EJ, Kittles RA, Argyropoulos C, Pfaff CL, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey WT, Jin P, McKeigue PM, Kamboh MI, Ferrell RE, Pollitzer WS, Shriver MD. 2001. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. Am J Phys Anthropol 114:18–29.

Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL. 2005. Support for the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proc Natl Acad Sci USA 102:15942–15947.

Relethford JH. 1994. Craniometric variation among modern human populations. Am J Phys Anthropol 95:53–62.

Relethford JH. 1997. Hemispheric difference in human skin color. Am J Phys Anthropol 104:449–457.

Relethford JH. 2000. Human skin color diversity is highest in sub-Saharan African populations. Hum Biol 72:773–780.

Relethford JH. 2002. Apportionment of global human genetic diversity based on craniometrics and skin color. Am J Phys Anthropol 73:629–636.

Relethford JH. 2004a. Boas and beyond: migration and craniometric variation. Am J Hum Biol 16:379–386.

Relethford JH. 2004b. Global patterns of isolation by distance based on genetic and morphological data. Hum Biol 76:499–513.

Relethford JH. 2008. The human species: an introduction to biological anthropology, 7th ed. New York: McGraw-Hill.

Relethford JH, Blangero J. 1990. Detection of differential gene flow from patterns of quantitative variation. Hum Biol 62:5–25.

Relethford JH, Crawford MH, Blangero JC. 1997. Genetic drift and gene flow in post-famine Ireland. Hum Biol 69:443–465.

Relethford JH, Harpending HC. 1994. Craniometric variation, genetic theory, and modern human origins. Am J Phys Anthropol 95:249–270.

Roberts DF, Kahlon DPS. 1976. Environmental correlations of skin colour. Ann Hum Biol 3:11–22.

Roseman CC. 2004. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. Proc Natl Acad Sci USA 101:12824–12829.

Roseman CC, Weaver TD. 2004. Multivariate apportionment of global human craniometric diversity. Am J Phys Anthropol 125:257–263.

Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW. 2002. Genetic structure of human populations. Science 298:2381–2385.

Sarich V, Miele F. 2004. Race: the reality of human differences. Boulder, CO: Westview.

Sauer NJ. 1992. Forensic anthropology and the concept of race: if races don't exist, why are forensic anthropologists so good at identifying them? Soc Sci Med 34:107–111.

Shanklin E. 1994. Anthropology and race. Belmont, CA: Wadsworth.

Templeton AR. 1998. Human races: a genetic and evolutionary perspective. Am Anthropol 100:632–650.

Templeton AR. 2007. Genetics and recent human evolution. Evolution 61:1507–1519.

Williams-Blangero S, Blangero J. 1989. Anthropometric variation and the genetic structure of the Jirels of Nepal. Hum Biol 61:1–12.

Williams-Blangero S, Blangero J. 1992. Quantitative genetic analysis of skin reflectance: a multivariate approach. Hum Biol 64:35–49.

Wilson EO, Brown WL Jr. 1953. The subspecies concept and its taxonomic application. Syst Zool 2:97–111.